



Tecnológico de Monterrey

Inteligencia artificial avanzada para la ciencia de datos I

Módulo 2
Grupo 101

Roberto Valdez Jasso

A01746863

**Módulo 2 Implementación de una técnica de aprendizaje máquina
con el uso de un framework**

Estado de México

Viernes 09 de septiembre de 2022.

Descripción:

Implementación de una técnica de aprendizaje maquina sin el uso de un framework.

- Crea un repositorio de GitHub para este proyecto.
- Programa uno de los algoritmos vistos en el módulo (o que tu profesor de módulo autorice) sin usar ninguna biblioteca o framework de aprendizaje máquina, ni de estadística avanzada. Lo que se busca es que implementes manualmente el algoritmo, no que importes un algoritmo ya implementado.
- Prueba tu implementación con un conjunto de datos y realiza algunas predicciones. Las predicciones las puedes correr en consola o las puedes implementar con una interfaz gráfica apoyándote en los vistos en otros módulos.
- Tu implementación debe de poder correr por separado solamente con un compilador, no debe depender de un IDE o de un “notebook”. Por ejemplo, si programas en Python, su implementación final se espera que esté en un archivo .py no en un Jupyter Notebook.
- Después de la entrega intermedia se te darán correcciones que puedes incluir en tu entrega final.

Actividad:

Tras haber realizado la actividad, puedo denotar lo siguiente:

En primera instancia puedo comentar que momento de generar el código de modelo de IA, llegué a entender y comprender la base de cómo funciona, es decir, el modelo de regresión logística realizada en Python 3.9 / Pycharm, entendí el cómo las funciones mínimas que debe tener para la realización de este. A comparación este modelo o generación de modelo por librería esta más sencillo ya que, en primera el modelo ya tiene carga consigo todas las herramientas, funciones y procesos necesarios para la generación y proceso de la regresión logística lo cual facilita el trabajo al momento de crearlo, como tambien los hiperparametros no se necesitan especificar, y se pueden calcular dentro de las mismas funciones. Ahora bien, no todo es fácil ya que hay que acomodar los datos como se nos pide en la regresión logística (binarios, falso o verdaderos, grupos o numéricos y no variables continuas) para la funcionalidad apta del modelo. Por otro lado, una vez corriendo este modelo a la par con el modelo manual de regresión logística de la actividad anterior y asi tambien compáralo con mi modelo generado, salieron los siguientes resultados:

Primera Prueba

Con Librería

```
Reporte de clasificacion :
```

	precision	recall	f1-score	support
0.0	0.82	0.88	0.85	109
1.0	0.79	0.70	0.74	69
accuracy			0.81	178
macro avg	0.80	0.79	0.79	178
weighted avg	0.81	0.81	0.81	178

Con Modelo Manual

```
Reporte de clasificacion sin crossvalidation :
```

	precision	recall	f1-score	support
0.0	0.64	0.99	0.77	109
1.0	0.88	0.10	0.18	69
accuracy			0.65	178
macro avg	0.76	0.55	0.48	178
weighted avg	0.73	0.65	0.54	178

```
/-----/  
Matrix de confusion de la predicion:  
[[377  63]  
 [ 91 180]]
```

Con estos resultados a la mano podemos decir lo siguiente:

El realizado con la librería tiene un mayor puntaje de precisión (81%) debido a sus constante conteo y revisión del error, lo cual hace que sus épocas sen considerablemente más exactas para asi satisfacer la revisión de la iteración, haciendo más preciso el proceso con los pasos realizados con sus ratio de aprendizaje que se va modificando de acuerdo al dataframe seleccionado o tambien por la interacción de los mismos.

Ahora bien, en el modelo generado por librería nos facilita en gran parte de los cálculos “tediosos” que se generan al momento de generar esta regresión, como el calculo de margen de error que se tiene en lo más mínimo posible para que los datos sigan iterando , para asi satisfacer la condición y encontrar las épocas exactas y necesarias para el dataset seleccionado, como tambien el ratio de aprendizaje, la cual, dependiendo por dataset de entrenamiento generado, este será calculado para satisfacer la cantidad de “pasos” necesarios y pesos del gradiente para tener una mayor precisión en el mismo, por ultimo para poder evaluar el modelo podemos usar un reporte de clasificación, a cual nos presentara el porcentaje de precisión con la cantidad de datos que fueron acertados, como tambien realizar una matriz de confusión para mayor entendimiento de que filas fueron exitosas y cuales no fueron exitosas al momento de predecir con sus validaciones y comprender asi a profundidad la precisión del mismo.